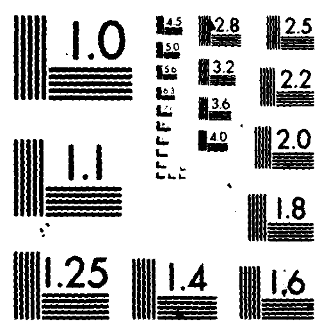
AD-A093 187 INDIANA UNIV AT BLOOMINGTON DEPT OF MATHEMATICS F/G 12/1
SHORTED MATRICES - AN EXTENDED CONCEPT AND SOME APPLICATIONS IN--ETC(
1980 S K MITRA, M L PURI AFOSR-76-2927

UNCLASSIFIED AFOSR-TR-80-1294 NL

END
DATE
FILMED
1 - 81
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 80-1294

(4)

SHORTED MATRICES - AN EXTENDED CONCEPT AND SOME APPLICATIONS IN DESIGN AND ANALYSIS OF EXPERIMENTS

Sujit Kumar Mitra (1)

Indiana University
Bloomington, Indiana
and

Indian Statistical Institute
New Delhi, India

and

Madan L. Puri(2)

Indiana University
Bloomington, Indiana

1. INTRODUCTION

Anderson and Duffin [2] have introduced the concept of "parallel sum" of a pair of matrices and have deduced interesting and important properties of this operation when the matrices concerned are non-negative definite. They were led to this concept from a parallel connection of "resistors" through a vectorial generalization of Kirchoff's and Ohm's laws in which resistors become non-negative definite linear operators. The concept of parallel sum was extended and its elegance further demonstrated by Rao and Mitra [19] who showed that most of the properties proved by Anderson and Duffin [2]

---

(1) Research supported by National Science Foundation Grant No. MCS76-00951 at Indiana University.

(2) Research supported by the Air Force Office of Scientific Research, AFSC, USAF under Grant No. AFOSR-76-2927. Reproduction in whole or in part permitted for any purpose of the United States Government.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR-80-1294	2. GOVT ACCESSION NO.: AD-A093 187
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): SHORTED MATRICES - AN EXTENDED CONCEPT AND SOME APPLICATIONS IN DESIGN AND ANALYSIS OF EXPERIMENTS	5. TYPE OF REPORT & PERIOD COVERED: Interim
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Sujit Kumar Mitra and Madan L. Puri	8. CONTRACT OR GRANT NUMBER(s): AFOSR-76-2927; NSF ...-76-...
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Indiana University, Department of Mathematics, Bloomington, Indiana 47405	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM (AFSC), Bolling Air Force Base, D.C. 20332	12. REPORT DATE: 1980
	13. NUMBER OF PAGES: 29
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Shorted matrices; parallel sum; linear operators, parallel summable interblock information, incomplete block experiments, linear models.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The concept of a shorted operator on the cone C_n of nonnegative definite matrices of order $n \times n$ introduced by Krein and studied recently by Anderson and Trapp ~~(Siam Journ. Appl. Math. 1975)~~ is extended to a wider class of matrices. For matrices in C_n, the shorting operation is permissible with reference to any subspace

(over)

DD FORM 1 JAN 73 1473

410522

UNCLASSIFIED

20. (continued)

S of the n dimensional Euclidean space E^n, provided restrictions are symmetrically placed on the row and column spans. For general matrices, the shorting operation is uniquely defined if and only if certain conditions are satisfied by the matrix itself and by subspaces providing restrictions on the row and column spans. The key point in this development is a theorem of Anderson and Trapp which exhibits the shorted n.n.d. matrix as the limit of a sequence of parallel sum operators. Some applications of the shorted operator in mathematical statistics with special reference to the design and analysis of experiments are provided.

/ 4, the nth power

are indeed true for a much wider class of pairs of matrices, designated by these authors as "parallel summable". Similar extensions of the concept of "hybrid sum", introduced by Duffin and Trapp [8] in analogy with a hybrid connection of resistors, were made by Mitra and Trapp [17]. The object of this paper is to offer a comparable extension of the notion of a shorted operator studied by Anderson [1] (3), by Anderson and Trapp [3] and the present authors [14]. The key point in this development is a theorem of Anderson and Trapp which exhibits the shorted n.n.d. matrix as the limit of a sequence of parallel sum matrices. We also describe some applications of the shorted operator in mathematical statistics.

2. PRELIMINARIES

Before we proceed any further let us record here some known results on the parallel sum. We follow the same notation as in our earlier paper [14].

Definition 2.1. Matrices A and B in $C^{m\times n}$ are said to be parallel summable (p.s.[19]) if $A(A+B)^{-}B$ is invariant under the choice of the generalized inverse $(A+B)^{-}$. If A and B are p.s., $P(A,B) = A(A+B)^{-}B$ is called the parallel sum of A and B.

The following theorem is proved in [19].

---

(3) See also Krein [10].

Theorem 2.1. A and B are p.s. iff

$$M(A) \subset M(A + B) \ , \ M(A^*) \subset M(A^* + B^*)$$

or equivalently

$$M(B) \subset M(A + B) \ , \ M(B^*) \subset M(A^* + B^*)$$

Theorem 2.2 lists certain known properties of the parallel sum [2,13,19].

Theorem 2.2. If A and B are p.s. matrices in $C^{m\times n}$, then

(a) $P(A,B) = P(B,A)$,

(b) A^* and B^* are also p.s. and $P(A^*,B^*) = [P(A,B)]^*$,

(c) $P(A,B)$ is n.n.d. when $m = n$ and A,B are n.n.d.,

(d) for C of rank m in $C^{p\times m}$, CA and CB are p.s. and $P(CA,CB) = CP(A,B)$

(e) $\{[P(A,B)]^-\} = \{A^- + B^-\}$

(f) $M[P(A,B)] = M(A) \cap M(B)$.

(g) if P_* is the orthogonal projector onto $M(A^*) \cap M(B^*)$ and P is the orthogonal projector onto $M(A) \cap M(B)$, then $[P(A,B)]^+ = P_*(A^- + B^-)P$.

(h) $P[P(A,B),C] = P[A,P(B,C)]$ if all the parallel sum operations are defined.

(i) if P_A and P_B are the orthogonal projectors onto $M(A)$ and $M(B)$, respectively, then the orthogonal projector onto $M(A) \cap M(B)$ is given by $P = 2P(P_A,P_B)$

Definition 2.2. If S is a subspace of E^m and $A \in C_m$, the shorted matrix $\mathcal{S}(A)$ is the unique matrix in $\mathcal{C}_m$ such that

$$\mathcal{M}[\mathcal{S}(A)] \subset \mathcal{S}$$

$$A \geq \mathcal{S}(A)$$

if $C \in \mathcal{C}_m$, $A \geq C$ and $\mathcal{M}(C) \subset \mathcal{S}$ then $\mathcal{S}(A) \geq C$.

The existence of $\mathcal{S}(A)$ was established by Anderson and Trapp[3] .

Theorem 2.3. Let $A,B \in \mathcal{C}_m$ and $\mathcal{M}(B) = \mathcal{S}$, then

$$\mathcal{S}(A) = \lim_{\lambda\to 0} \frac{1}{\lambda} P(\lambda A,B)$$

Theorem 2.3 was proved by Anderson and Trapp for the special case where B is the orthogonal projector (Theorem 12 [3]). The general case could be proved on same lines. Alternatively, a direct proof could be constructed using simultaneous diagonalization of the pair A,B of n.n.d. matrices (see e.g. Theorem 6.2.3,[19]).

3. The Shorted Matrix - An Extended Concept

Let the pair of matrices A,B in $\mathcal{C}^{m\times n}$ be p.s. and

$$\lim_{\lambda\to 0} A(\lambda A + B)^{-} B = C \tag{3.1}$$

exist and be finite.

Theorem 3.1 gives certain properties of the C matrix when it exists.

Theorem 3.1.

(a) $C = \lim_{\lambda \to 0} B(\lambda A + B)^{-}A$ (3.2)

(b) $M(C) \subset M(A) \cap M(B)$

$M(C^*) \subset M(A^*) \cap M(B^*)$

(c) $M(A - C)$ is virtually disjoint with $M(B)$ and so is $M(A^* - C^*)$ with $M(B^*)$

(d) C and $A - C$ are disjoint matrices [12], that is

$$M(A) = M(C) \oplus M(A-C)$$

$$M(A^*) = M(C^*) \oplus M(A^* - C^*) \quad (3.3)$$

(e) $$M(C) = M(A) \cap M(B)$$

$$M(C^*) = M(A^*) \cap M(B^*) \quad (3.4)$$

(f) Let E be any matrix such that $M(E) \subset M(B)$, $M(E^*) \subset M(B^*)$, then

$$\text{Rank } (A - E) \geq \text{Rank } (A - C)$$

the sign of equality holding if and only if $E = C$.

Proof: (a). If A and B are p.s., λA and B are p.s. for each λ sufficiently small. Hence

$$C = \lim_{\lambda \to 0} \frac{1}{\lambda} P(\lambda A, B)$$

Since $P(\lambda A, B) = P(B, \lambda A)$ by Theorem 2.2(a), (3.2) follows.

(b). Consider a typical vector $C\underset{\sim}{x}$ in $M(C)$

$$C\underset{\sim}{x} = \lim_{\lambda \to 0} \underset{\sim}{y}_\lambda$$

Where $\underset{\sim}{y}_\lambda = A(\lambda A + B)^{-}Bx \in M(A)$. Since $M(A)$ is closed,

$$C\underset{\sim}{x} = \lim_{\lambda \to 0} \underset{\sim}{y}_\lambda \in M(A).$$

Similarly using (3.2) we have $C\underset{\sim}{x} \in M(B)$ and hence $C\underset{\sim}{x} \in M(A) \cap M(B)$. The other part of (b) is established in a like manner.

(c). Let $(A - C)\underset{\sim}{u} = B\underset{\sim}{v}$ be a vector in $M(A - C) \cap M(B)$. Then $A\underset{\sim}{u} = C\underset{\sim}{u} + B\underset{\sim}{v}$.

Hence $$C\underset{\sim}{u} = \lim_{\lambda\to 0} B(\lambda A + B)^{-}A\underset{\sim}{u} = \lim_{\lambda\to 0} B(\lambda A + B)^{-}(C\underset{\sim}{u} + B\underset{\sim}{v})$$

$$\Longrightarrow \lim_{\lambda\to 0} B(\lambda A + B)^{-}B\underset{\sim}{v} = \lim_{\lambda\to 0} \lambda A(\lambda A + B)^{-}C\underset{\sim}{u} \tag{3.5}$$

Since $C = BK$ for some matrix K in $C^{n\times m}$, the RHS of (3.5) is seen to be equal to the null vector, while

$$B = B(\lambda A + B)^{-}(\lambda A + B)$$

for each λ sufficiently small implies on taking limits of both sides that

$$B = \lim_{\lambda\to 0} B(\lambda A + B)^{-}B$$

Hence $B\underset{\sim}{v} = \underset{\sim}{0}$ and the first part of (c) is established. The proof of the second part is similar.

(d). (d) is a simple consequence of (c).

(e). Let $\underset{\sim}{x} \in M(A) \cap M(B)$. Using (3.3) we write

$$\underset{\sim}{x} = \underset{\sim}{x}_1 + \underset{\sim}{x}_2$$

where $\underset{\sim}{x}_1 \in M(C)$ and $\underset{\sim}{x}_2 \in M(A - C)$. Observe that $\underset{\sim}{x}_2 = \underset{\sim}{x} - \underset{\sim}{x}_1 \in M(B)$ and hence $\in M(B) \cap M(A - C)$. This implies $\underset{\sim}{x}_2 = \underset{\sim}{0}$ and $\underset{\sim}{x} = \underset{\sim}{x}_1 \in M(C)$. This, in view of (b) establishes the first part of (e). The other part is similarly deduced.

(f) In view of (c), the expression

$$A - E = (C - E) + (A - C)$$

exhibits $A - E$ as the sum of two disjoint matrices $C - E$ and $A - C$. Hence

$$\text{Rank}\ (A - E) = \text{Rank}\ (C - E) + \text{Rank}\ (A - C)$$

This concludes the proof of (f) and of Theorem 3.1.

The matrix C will henceforth be called the matrix A shorted by the matrix B and denoted by $S(A|B)$. Theorem 3.2 gives some more properties of the shorted matrix.

Theorem 3.2. Let $S(A|B)$ exist. Then

(a) $S(A^*|B^*)$ also exists and $S(A^*|B^*) = [S(A|B)]^*$,

(b) if $K \in C^{p\times m}$ and Rank $K = m$, $S(KA|KB)$ exists and $S(KA|KB) = KS(A|B)$

(c) if $m = n$, A is n.n.d. and further

$$\mathcal{M}(B) \cap \mathcal{M}(A) = \mathcal{M}(B^*) \cap \mathcal{M}(A) \tag{3.6}$$

$S(A|B)$ is n.n.d.

Proof: Proofs of (a) and (b) are straightforward and are omitted. (c). Observe that by the (a) part if $C = S(A|B)$, then $C^* = S(A|B^*)$. Further, if (3.6) holds both C and C^* have identical row and column spans. Application of theorem 3.1(f) now shows $C = C^*$.

Choose and fix a hermitian g-inverse A^- of A. If for some $\underset{\sim}{x}$, $\underset{\sim}{x}^* C \underset{\sim}{x} < 0$,

$$\underset{\sim}{x}^*CA^-AA^-C\underset{\sim}{x} = \underset{\sim}{x}^* C \underset{\sim}{x} < 0 .$$

The equality follows from the fact that A^-AA^- is a g-inverse of

A and A being the sum of disjoint matrices C and $A - C$ every g-inverse of A is a g-inverse of C. This contradicts the assumption that A is n.n.d., and establishes the claim in (c).

That (c) is not true in general can be seen from the following counter example. Let I and H be the identity matrix and an idempotent matrix in $C^{m\times m}$. It is not difficult to see that $S(I|H)$ exists and is equal to H. H need not be hermitian, and it is conceivable that a vector $\underset{\sim}{x}$ might exist such that $\underset{\sim}{x}^* H \underset{\sim}{x} < 0$. Consider the idempotent matrix

$$H = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

for such a counter example.

Remarks. A matrix $A \in C^{m\times m}$ is said to be almost positive definite (a.p.d.) [7,11] if $\forall\ x \in E^m$, $\mathrm{Re}\,(x^* A x) \geq 0$ and $x^* A x = 0 \Rightarrow Ax = 0$. Unlike n.n.d. matrices, an a.p.d. matrix need not be hermitian. Similar to theorem 3.2(c), we can prove that if A is a.p.d., $S(A|B)$ exists and (3.6) holds, then $S(A|B)$ is a.p.d.

This can be proved as follows. Since A is a.p.d., it follows as in Theorem 2 [11] that A is an EP matrix (that is $\mathcal{M}(A) = \mathcal{M}(A^*)$). Equations (3.4) and (3.6) therefore imply that $C = S(A|B)$ is EP. Further as in Corollaries 2 and 3 of [11], A^+ is seen to be a.p.d. Since C is EP

$$C[I - A^+C] = 0 \Rightarrow C^*[I - A^+C] = 0 \Rightarrow C^* = C^*A^+C$$

If $\mathrm{Re}\,(x^*Cx) < 0$, then $\mathrm{Re}(x^*C^*x) = \mathrm{Re}(x^*C^*A^+Cx) < 0$, which contradicts almost positive definiteness of A^+ . Also,

$$x^* C x = 0 \Rightarrow A^+ C x = 0 \Rightarrow Cx = 0 .$$

A matrix $A \in \mathbb{C}^{m\times m}$ is said to be positive semidefinite (p.s.d.) if $\forall x \in \mathbb{C}^m$, $\mathrm{Re}(x^* A x) \geq 0$. When A is p.s.d. and (3.6) holds, $S(A|B)$ if it exists is also p.s.d. This can be proved on similar lines.

Let the matrices A and B be p.s. and $A + B$ be of rank r . Consider a rank factorization of $A + B$:

$$A + B = LR$$

and the representations

$$A = LDR , \; B = L(I - D) R \tag{3.7}$$

implied by parallel summability, where D and $I - D$ are square matrices in $\mathbb{C}^{r\times r}$. It is easily seen that in any such representation, the matrix D (and naturally $I - D$) is uniquely determined up to a similarity transformation. Theorem 3.3 gives a necessary and sufficient condition for $S(A|B)$ to exist.

Theorem 3.3. Let A and B be p.s., $S(A|B)$ exists iff $I - D$ is of Drazin index 1, that is

$$\mathrm{Rank}\ (I - D)^2 = \mathrm{Rank}\ (I - D) \tag{3.8}$$

or equivalently

$$\mathrm{Rank}\ B(A + B)^- B = \mathrm{Rank}\ B \tag{3.9}$$

Proof: Without any loss of generality we assume that D is already in Jordan canonical form, and write D as the sum of two disjoint matrices D_1 and D_2, each of order $r \times r$ where

D_1 is identical with D everywhere including in all its diagonal Jordan blocks, except for the Jordan block corresponding to the eigen value 1, if any, which is replaced by a null matrix and $D_2 = D - D_1$. The above assumption (3.8) implies that $(I - D)$ is disjoint with D_2. Hence

$$\begin{aligned} A(\lambda A + B)^{-}B &= LD(\lambda D + I - D)^{-}(I - D)R \\ &= L(D_1 + D_2)[\lambda D_1 + (I - D) + \lambda D_2]^{-}(I - D)R \\ &= LD_1\ [\lambda D_1 + (I - D) + \lambda D_2]^{-}(I - D)R \\ &= LD_1[\lambda D_1 + (I - D)]^{-}(I - D)R \\ &= LD_1R - \lambda LD_1[\lambda D_1 + (I - D)]^{-}D_1R \end{aligned}$$

Taking limit as $\lambda \to 0$ we have

$$\lim_{\lambda \to 0} A(\lambda A + B)^{-}B = LD_1R$$

This concludes the proof of the 'if' part.

Assume now that (3.1) holds and consider the representation

$$A + B = (B + C) + (A - C)$$

where $B + C$ and $A - C$ are disjoint matrices. Since A and B are p.s., $B + C$ has the same row and column spans as that of B. Consider rank factorizations of $B + C$ and $A - C$:

$$B + C = L_1R_1$$

$$A - C = L_2R_2$$

leading to the rank factorization $A + B = LR$ where

$$L = (L_1 : L_2) \quad \text{and} \quad R = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix}$$

If $B = L_1FR_1$, clearly F is nonsingular. Hence $I - D = \mathrm{diag}(F,0)$

is of Drazin index 1. The last part of Theorem 3.3 is trivial. This concludes proof of Theorem 3.3.

Theorem 3.4. If $S(A|B)$ is defined, then $S[A|S(A|B)]$ is also defined and

$$S[A|S(A|B)] = S(A|B) \tag{3.10}$$

Proof: The proof is fairly straightforward and is omitted.

Theorem 3.5. A general solution to a g-inverse of $S(A|B)$ is $A^- + X_b$ where A^- is an arbitrary g-inverse of A and X_b is any arbitrary solution of the homogeneous equation

$$BX_bB = 0 \tag{3.11}$$

Proof: Every g-inverse of A is a g-inverse of $S(A|B)$. This we have already noted while proving theorem 3.2. Consider now a matrix $A^- + X_b$ as determined above

$$S(A|B)(A^- + X_b)\ S(A|B) = S(A|B) + S(A|B)X_bS(A|B) = S(A|B)$$

on account of Theorem 3.1(b). This shows

$$\{A^- + X_b\} \subset \{[S(A|B)]^-\}$$

Now choose and fix a g-inverse G of A . A general solution to a g-inverse of $S(A|B)$ is $G + X_s$ where X_s is a general solution to the homogeneous equation

$$S(A|B)X_sS(A|B) = 0 \text{ .}$$

Theorem 3.1(e) implies that any such matrix X_s can be written as $X_s = X_a + X_b$ where X_a and X_b satisfy respectively the equations

$$AX_aA = 0 \ , \ BX_bB = 0 \ .$$

The matrix $G + X_a \in \{A^-\}$. Hence

$$\{[S(A|B)^-\} \subset \{A^- + X_b\}$$

Theorem 3.5 is thus proved.

Theorem 3.6. (a) If $S(A|B)$ and $S(B|A)$ are both defined, then $S(A|B)$ and $S(B|A)$ are p.s. and

$$P[S(A|B), S(B|A)] = P(A,B) \tag{3.12}$$

(b) $$\{[S(A|B)]^- + [S(B|A)]^-\} = \{A^- + B^-\} \tag{3.13}$$

Proof: (a). From theorem 3.3 it is seen that $S(B|A)$ will exist iff Rank D^2 = Rank D . Let $(I - D)_1$ denote a matrix which is identical with $I - D$ everywhere except for the diagonal block corresponding to the zero eigen value of D , which is replaced by a null matrix. As in the proof of Theorem 3.3, it is seen that

$$S(B|A) = \lim_{\lambda \to 0} A(\lambda B + A)^- B = L(I - D)_1 R$$

When both $S(A|B)$ and $S(B|A)$ exist it is seen that D_1 and $(I - D)_1$ which are both block diagonal, have nonnull diagonal blocks at identical positions, each such nonnull pair adding up to an identity matrix of same order as of the diagonal block concerned. This shows that $S(A|B)$ and $S(B|A)$ are parallel summable and

$$P[S(A|B),S(B|A)] = LD_1(I - D)_1R = LD(I - D)R$$
$$= P(A,B)$$

(b). Clearly $\{A^- + B^-\} \subset \{[S(A|B)]^- + [S(B|A)]^-\}$.
Conversely by theorem 3.5, $[S(A|B)]^-$ can be written as $G_a + X_b$ where $G_a \in \{A^-\}$ and $BX_bB = 0$. Similarly $[S(B|A)]^-$ can be written as $G_b + X_a$ where $G_b \in \{B^-\}$ and $AX_aA = 0$. Hence

$$[S(A|B)]^- + [S(B|A)]^- = G_a + X_b + G_b + X_a$$
$$= (G_a + X_a) + (G_b + X_b) \; .$$

This shows $\{[S(A|B)]^- + [S(B|A)]^-\} \subset \{A^- + B^-\}$ which concludes the proof of the part (b) and of theorem 3.6.

Theorem 3.7.

$$S[P(A,B)|C] = P[S(A|C),B] = P[S(B|C),A] \tag{3.14}$$

when the parallel sum and shorted matrices involved are defined.

Proof. Using Theorem 3.5 and Theorem 2.2(e)

$$\{(S[P(A,B)|C])^-\} = \{A^- + B^- + X_c\}$$
$$\{(P[S(A|C),B])^-\} = \{A^- + X_c + B^-\}$$
$$\{(P[S(B|C),A])^-\} = \{B^- + X_c + A^-\}$$

This shows the three matrices in (3.14) have identical general solutions for a g-inverse. Since a matrix is uniquely determined by its class of g-inverses (Theorem 2.4.2, [19]) (3.14) is established.

The following theorem can also be proved using a similar argument. We omit the proof.

Theorem 3.8.

$$\begin{aligned} \mathcal{S}[\mathcal{S}(A|B)|C] &= \mathcal{S}[A|\mathcal{S}(B|C)] = \mathcal{S}[A|\mathcal{S}(C|B)] \\ &= \mathcal{S}(A|P(B,C)] \end{aligned} \tag{3.15}$$

when the parallel sum and shorted matrices involved are defined.

4. Another Approach

In view of Theorem 3.1(f), one is tempted to put forward the following definition of the shorted matrix imitating the Anderson-Trapp definition given in Section 2.

Let A be a given matrix in $C^{m\times n}$ and $\mathcal{S},\mathcal{T}$ be given subspaces in E^m, E^n respectively. The shorted matrix $S(A|\mathcal{S},\mathcal{T})$ is a matrix C in $C^{m\times n}$, such that

$$\text{(a)} \quad \mathcal{M}(C) \subset \mathcal{S}\ ,\ \mathcal{M}(C^*) \subset \mathcal{T} \tag{4.1}$$

$$\begin{aligned} \text{(B)} \quad & \text{if } E \in C^{m\times n},\ \mathcal{M}(E) \subset \mathcal{S},\ \text{and } \mathcal{M}(E^*) \subset \mathcal{T}, \\ & \text{then } \operatorname{Rank}(A - E) \geq \operatorname{Rank}(A - C) \end{aligned} \tag{4.2}$$

This definition however does not always lead to a unique answer. Consider for example the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

and let

$$\mathcal{S} = \mathcal{T} = \mathcal{M}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

It is seen that for arbitrary scalars a and b the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ a & b & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

satisfies the conditions required of the matrix C in the above definition. The following theorem gives necessary and sufficient conditions on the triplet (A,S,T) so that $S(A|S,T)$ may exist uniquely.

Let S and T be column spans of matrices L_1 and $R_1{}^*$ respectively, and let the columns of L_2 and $R_2{}^*$ span respectively complementary subspaces of S in E^m and of T in E^n. We assume that $L_1, L_2, R_1{}^*, R_2{}^*$ are of full column ranks. Let us write

$$A = (L_1 : L_2) \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \tag{4.3}$$

Theorem 4.1. The shorted matrix $S(A|S,T)$ exists and is unique iff

$$M(W_{21}) \subset M(W_{22}) \ , \ M(W_{12}) \subset M(W_{22}^*) \tag{4.4}$$

When (4.4) is satisfied

$$S(A|S,T) = L_1(W_{11} - W_{12}W_{22}^{-}W_{21})R_1$$

Proof: The required conditions are seen to be independent of the specific choice of matrices L_1, L_2, R_1 and R_2 in the sense that if the conditions are met for one choice, they would also be met for an alternative choice.

Assume now that (4.4) holds and write

$$\begin{aligned} A &= L_1W_{11}R_1 + L_1W_{12}R_2 + L_2W_{21}R_1 + L_2W_{22}R_2 \\ &= L_1(W_{11} - W_{12}W_{22}^-W_{21})R_1 + (L_1W_{12} + L_2W_{22})(R_2 + W_{22}^-W_{21}R_1) \\ &= L_1(W_{11} - W_{12}W_{22}^-W_{21})R_1 + (L_1W_{12}W_{22}^- + L_2)(W_{21}R_1 + W_{22}R_2) \\ &= A_1 + A_2 \quad \text{, say} \end{aligned} \tag{4.5}$$

Clearly $\mathcal{M}(A_1) \subset S$, $\mathcal{M}(A_1^*) \subset T$. We shall show that $\mathcal{M}(A_2)$ is virtually disjoint with S and $\mathcal{M}(A_2^*)$ with T . Let

$$\begin{aligned} A_2\underset{\sim}{x} &= (L_1W_{12} + L_2W_{22})(R_2 + W_{22}^-W_{21}R_1)\underset{\sim}{x} \\ &= (L_1W_{12} + L_2W_{22})\underset{\sim}{y} \end{aligned}$$

be a vector in $\mathcal{M}(A_2) \cap \mathcal{M}(L_1)$. Then $L_{22}W_{22}\underset{\sim}{y} = \underset{\sim}{0}$

$$\Rightarrow W_{22}\underset{\sim}{y} = \underset{\sim}{0} \Rightarrow W_{12}\underset{\sim}{y} = \underset{\sim}{0} \text{ on account of (4.4)}$$

$$\Rightarrow A_2\underset{\sim}{x} = \underset{\sim}{0} ,$$

That $\mathcal{M}(A_2^*)$ is virtually disjoint with T is similarly established. That the matrix A_1 satisfies also the condition (4.2) required of the C matrix and is the unique matrix to do so follows as in the proof of theorem 3.1(f).

Conversely suppose that there exists a unique matrix C satisfying (4.1) and (4.2). This implies that $\mathcal{M}(A - C)$ is virtually disjoint with S and $\mathcal{M}(A^* - C^*)$ with T . Let $C = L_{11}R_{11}$, $A - C = L_{21}R_{21}$ be rank factorizations of C and $A - C$ respectively. Let the columns of $L_1 = (L_{11}:L_{12})$ provide a basis for S and those of $R_1^* = (R_{11}^*:R_{12}^*)$ a basis for T . Further let the columns of $(L_1:L_{21}:L_{22})$ form a basis of E^m and those of $(R_1^*:R_{21}^*:R_{22}^*)$ a basis for E^n . Let us write L_2 for the matrix

$(L_{21}:L_{22})$ and R_2 for the matrix $(R_{21}:R_{22})$. Clearly by construction the matrices C and $A - C$ are of the form $C = L_1W_{11}R_1$, $A - C = L_2W_{22}R_2$ which shows that in a representation ofthe type (4.3) , $W_{12} = 0$, $W_{21} = 0$ and the condition (4.4) is trivially satisfied.

In [17], Mitra and Trapp defined the generalized shorted operator as the strong hybrid sum of A with the null matrix. Theorem 4.1 is closely related with this definition. The reader is also referred to Theorem 2 in Carlson [5] which considers decompositions of the matrix A with a somewhat different emphasis.

When A is a square matrix, that is $m = n$ and $S = T$, condition (4.4) is seen to be equivalent to the condition that A be $S^{\perp}$ complementable (Ando [4]) where $S^{\perp}$ denotes the orthogonal complement of S . Further

$$S(A|S,S) = A_{/S^{\perp}}$$

the generalized Schur complement. The verification is fairly straightforward. For the case $m \neq n$, the notion could be extended as follows:

Let M , N be given subspaces in E^m, E^n respectively and following Ando's notation let I_M and I_N denote respectively the orthogonal projectors onto M and N under the respective dot products.

Definition. The matrix A is said to be M,N complementable if there exists matrices $M_\ell \in C^{m\times m}$, $N_r \in C^{n\times n}$ such that

$$M_\ell I_M = M_\ell \ , \ I_N N_r = N_r \tag{4.6}$$

$$I_M A N_r = I_M A \ , \ M_\ell A I_N = A I_N \ . \tag{4.7}$$

When (4.6) and (4.7) are satisfied, we have

$$M_\ell A = M_\ell A N_r = A N_r \tag{4.8}$$

$M_\ell A N_r$ which clearly depends only on M,N and A is called the Schur compression of A and denoted by $A_{M,N}$. $A - M_\ell A N_r$ is called the generalized Schur complement of A and denoted by $A_{/M,N}$.

Theorem 4.2. Let A be M,N complementable. Then

$$A_{/M,N} = S(A|M^\perp,N^\perp) \tag{4.9}$$

Proof: Observe that

$$I_M(A - M_\ell A N_r) = I_M(A - A N_r) = 0$$

$$(A - M_\ell A N_r) I_N = (A - M_\ell A) I_N = 0 \ .$$

These show that

$$M(A_{/M,N}) \subset M^\perp \ , \ M[(A_{/M,N})^*] \subset N^\perp$$

Further if

$$M_\ell A N_r a = A N_r a \in M(A_{M,N}) \cap M^\perp$$

$A N_r a = I_M A N_r a = 0$. This shows $M(A_{M,N})$ is virtually disjoint with $M^\perp$. Similarly $M[(A_{M,N})^*]$ is seen to be virtually disjoint with $N^{\perp}$. Hence $A_{/M,N}$ is seen to be the unique matrix satisfying (4.1) and (4.2) with $S = M^{\perp}$, $T = N^{\perp}$.

For an application of Theorem 4.1 consider the following problem. Let $A,B \in C^{m\times n}$ and Λ denote the matrix

$$\Lambda = \begin{pmatrix} A & A \\ A & A+B \end{pmatrix}$$

Put

$$S = M\begin{pmatrix} I_m \\ 0 \end{pmatrix} \quad , \quad T = M\begin{pmatrix} I_n \\ 0 \end{pmatrix}$$

By Theorem 4.1, $S(\Lambda|S,T)$ exists uniquely iff $M(A) \subset M(A+B)$, $M(A^*) \subset M(A^*+B^*)$, that is, if A and B are p.s. Further, if this condition is satisfied

$$S(\Lambda|S,T) = \begin{pmatrix} A - A(A+B)^{-}A & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} P(A,B) & 0 \\ 0 & 0 \end{pmatrix} \quad (4.10)$$

Anderson and Trapp [3] have used (4.10) to define the parallel sum of hermitian n.n.d. matrices through the concept of a shorted operator.

5. Computation of the Shorted Matrix

Let $A \in C^{m\times n}$, $X \in C^{m\times p}$, $Y \in C^{q\times n}$. Let F denote the matrix $\begin{pmatrix} A & X \\ Y & 0 \end{pmatrix}$ where 0 is the null matrix in $C^{q\times p}$. The following result can easily be established. We omit the proof.

Lemma 5.1. Condition (4.4) is equivalent to each of the following conditions.

(a) $\quad \text{Rank } F = \text{Rank } (A\ X) + \text{Rank } Y = \text{Rank} \begin{pmatrix} A \\ Y \end{pmatrix} + \text{Rank } X \qquad (5.1)$

(b) $\quad M\begin{pmatrix} A \\ 0 \end{pmatrix} \subset M(F)\ ,\ M\begin{pmatrix} A^* \\ 0 \end{pmatrix} \subset M(F^*) \qquad (5.2)$

Let $G = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix} \in \{F^-\}$.

Lemma 5.2. If condition (5.1) or equivalent (5.2) holds, then

(i) $\quad XC_3X = X\ ,\ YC_2Y = Y$

(ii) $\quad YC_1X = 0\ ,\ AC_1X = 0\ ,\ YC_1A = 0$

(iii) $\quad AC_2Y = XC_3A = XC_4Y$

(iv) $\quad AC_1AC_1A = AC_1A\ ,\ \text{Tr } AC_1 = \text{Rank } (A \vdots X) - \text{Rank } X$
$\qquad = \text{Rank} \begin{pmatrix} A \\ Y \end{pmatrix} = \text{Rank } Y$

(v) $\quad \begin{pmatrix} C_1 \\ C_3 \end{pmatrix} \in \{(A \vdots X)^-\}\ ,\ (C_1 \vdots C_2) \in \left\{ \begin{pmatrix} A \\ Y \end{pmatrix}^- \right\}$

Proof: Lemma 5.2 except for the second part of (iv) follows from the following equations

$$FGF = F \tag{5.3a}$$

$$FG\begin{pmatrix} A \\ 0 \end{pmatrix} = \begin{pmatrix} A \\ 0 \end{pmatrix} \tag{5.3b}$$

$$(A\vdots 0)GF = (A\vdots 0) \tag{5.3c}$$

the last two equations being consequences of (5.2).

Since $G \in \{F^-\}$, $C_3 \in \{X^-\}$, $C_2 \in \{Y^-\}$,

$$\begin{aligned}\text{Rank } F = \text{Rank } FG = \text{Tr } FG &= \text{Tr } (AC_1 + XC_3) + \text{Tr } YC_2 \\ &= \text{Tr } AC_1 + \text{Rank } X + \text{Rank } Y .\end{aligned}$$

The second part of (iv) therefore follows from (5.1)

Theorem 5.1. Let $S = M(X)$, $T = M(Y^*)$. If condition (5.1) or equivalently (5.2) holds, the matrix

$$XC_4Y = AC_2Y = XC_3A \tag{5.4}$$

is the unique shorted matrix $(A|S,T)$.

Proof: (5.3b) $\Rightarrow$

$$AC_1A + XC_3A = A \tag{5.5}$$

Since $A - XC_3A = AC_1A$, it suffices to show that

$$M(AC_1A) \cap M(X) = \{0\} \tag{5.6a}$$

$$M[(AC_1A)^*] \cap M(Y^*) = \{0\} \tag{5.6b}$$

If $AC_1Aa = Xb \in M(AC_1A) \cap M(X)$, then

$$AC_1Aa = AC_1AC_1Aa = AC_1Xb = 0 \Rightarrow (5.6a).$$

(5.6b) is similarly established. The rest of the proof of Theorem 5.1 is similar to the proof of the 'if' part of Theorem 4.1.

Remark 1. Lemma 5.2 for the special case where A is a real symmetric n.n.d. matrix and $Y = X^1$ was found by Rao [18]. That C_2 in such a case is a minimum A seminorm g-inverse of Y was shown in [19, Corollary 1, p. 47]. Using Theorem 2.1 of Mitra and Puri [14] it is seen that $AC_2Y = \mathcal{S}(A)$ where $\mathcal{S} = \mathcal{M}(X)$. It is remarkable that in the general case, the same formula also provides the shorted matrix $S(A|\mathcal{S},\mathcal{T})$ when in no conceivable way, can C_2 be interpreted a minimum A seminorm g-inverse of Y.

Remark 2. Formula (5.4) appears to be direct and therefore simplier to compute compared to the expression given in Theorem 4.1, as it does not require the identification of complementary subspaces of $\mathcal{S}$ and $\mathcal{T}$ and determination of the W_{ij} matrices in (4.3).

Theorem 5.2. Let the matrices A and B in $\mathbb{C}^{m\times n}$ be p.s. and in addition let (3.8) or equivalently (3.9) hold, then the matrix $F = \begin{pmatrix} A & B \\ B & 0 \end{pmatrix}$ satisfies condition (5.1) and if $G = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix} \in \{F^-\}$, then $AC_2B = S(A|B)$.

Proof: $$\text{Rank } F = \text{Rank}\begin{pmatrix} A+B & B \\ B & 0 \end{pmatrix} = \text{Rank}\begin{pmatrix} A+B & B \\ 0 & -B(A+B)^-B \end{pmatrix}$$

since $(B^*) \subset (A^* + B^*)$.

The RHS is further equal to

$$\text{Rank } (A + B) + \text{Rank } [B(A + B)^-B] = \text{Rank } (A:B) + \text{Rank } B$$

$$= \text{Rank}\begin{pmatrix} A \\ B \end{pmatrix} + \text{Rank } B .$$

The second part of Theorem 5.2 follows from Theorem 5.1.

6. Some Applications of the Shorted Matrix

a. Recovery of interblock information in incomplete block experiments [6,9]

Consider a pair of consistent linear equations

$$Ax = a \tag{6.1}$$

$$Bx = b \tag{6.2}$$

and the combined equation

$$(A + B)x = a + b \tag{6.3}$$

We shall assume that $M(A) \subset M(A + b)$ so that the equations (6.3) may be consistent whenever (6.1) and (6.2) are so. This condition is satisfied for example when A and B are p.s. matrices. The linear function p^*x assumes a unique value for every solution x of (6.1) iff

$$p \in M(A^*) \tag{6.4}$$

Among such linear functions we are interested in identifying those for which substitution of a solution of (6.3) or of (6.1) leads toidentical answers. Such problems crop up in the theory of recovery of interblock information in incomplete block experiments where (6.1) and (6.2) are respectively the normal equations for deriving intra- and inter block estimates and (6.3) is the normal equation for deriving combined intra- inter block estimates. When $S(A|B)$ exists, a neat answer is provided by

Theorem 6.1.

Theorem 6.1. If $S(A|B) = C$ exists, then

$$p^*(A + B)^-(a + b) = p^*A^-a \quad \forall a \in M(A) \;, \; b \in M(B) \qquad (6.5)$$

$$\text{iff} \qquad p \in M(A^* - C^*) \qquad (6.6)$$

Proof: 'if part' . We note first that $M(A^* - C^*) \subset M(A^*)$.
Let x_o satisfy (6.1)

$$\begin{aligned} &(A - C)(A + B)^-(a + b) \\ &= (A - C)(B + C + A - C)^-[(A + B)x_o + b - Bx_o] \\ &= (A - C)(B + C + A - C)^-(A + B)x_o \end{aligned}$$

since $(B + C)$ and $(A - C)$ are disjoint matrices and $b - Bx_o \in M(B) = M(B + C)$. The RHS further simplifies to

$$(A - C)(A + B)^-(A + B)x_o = (A - C)x_o$$

since $M(A^* - C^*) \subset M(A^*) \subset M(A^* + B^*)$.
'only if part'. (6.5) $\Rightarrow$

$$\begin{aligned} &p^*(A + B)^-b = 0 \quad \forall \; b \in M(B) \\ \Rightarrow \; &p^*(A + B)^-B = 0 \end{aligned} \qquad (6.7)$$

$$(6.4) \Rightarrow p^* = \lambda^*A \quad \text{for some} \quad \lambda \in E^m \qquad (6.8)$$

Substituting (6.8) in (6.7) we have

$$\lambda^*P(A,B) = 0 \qquad (6.9)$$

Since $M[P(A,B)] = M[S(A|B)]$, and $A^- \varepsilon \{C^-\}$, it follows that

$$(6.9) \Rightarrow \lambda^* C = 0$$

$$\Rightarrow \lambda^* = \mu^*[I - CA^-] \text{ for some } \mu \in E^m . \quad (6.10)$$

Hence $p^* = \lambda^* A = \mu^*[I - CA^-]A = \mu^*(A - C) \Leftrightarrow$ (6.6)

This concludes the proof of Theorem 6.1.

(b) Test of linear hypothesis in linear models

Let the random vector $\underset{\sim}{Y} \sim N_n(X\beta, \sigma^2 I)$ where β an m-tuple is an unknown parameter vector and $\sigma^2 > 0$ is also an unknown parameter. X is a known matrix.

Consider a hypothesis

$$H_0 : H\beta = b \quad (6.11)$$

We shall assume that the equation (6.11) is consistent. As otherwise the hypothesis could be rejected without any formal statistical test. It was shown in [15] that when $H\beta$ is not estimable, only the estimable part of (6.11) could be tested. To be more precise, let K be a matrix such that

$$M(H'K') = M(H') \cap M(X')$$

where ' on a matrix indicates its transpose. Then one could only test if

$$KH\beta = Kh \quad (6.12)$$

and deviations from (6.11) that do not result in deviations from (6.12) will go undetected. In the same paper the authors suggested an expression for K . One could alternatively use

$$K = X'X(H'H + X'X)^- H'$$

in view of Theorem 2.2(f).

We shall however recommend

$$K = CH^-_{m(C)}$$

where $C = X'X$ is the matrix of normal equations that provide least squares estimates for the parameters β and $H^-_{m(C)}$ is a minimum C seminorm g-inverse of H.

Observe that

$$KH = S(C)$$

is the shorted matrix C where $S = M(H')$ [12]. The shorted matrix is symmetric and

$$M(KH) = M(H'K') = M(H') \cap M(C) = M(H') \cap M(X')$$

Further the dispersion matrix $D(KH\hat{\beta})$ of the BLUE of $KH\beta$ is given by

$$\begin{aligned} D(KH\hat{\beta}) &= [S(C)]C^-[S(C)]\sigma^2 \\ &= S(C)\sigma^2 . \end{aligned}$$

Note that C^-/σ^2 is a g-inverse of this dispersion matrix. If $\underset{\sim}{u} = S(C)\hat{\underset{\sim}{\beta}} - \underset{\sim}{g}$ where $\underset{\sim}{g} = CH^-_{m(C)}h$, we have the following simple formula for computing the expression which appears in the numerator of the variance ratio test. We note in this connection that under H_0,

$$\underset{\sim}{u}'C^{-}\underset{\sim}{u}/\sigma^2 \sim \chi^2_v$$

where χ^2_v is a chi square random variable with the degrees of freedom $v = \text{Rank } S(C) = \text{trace } C^{-}(C)$. Since any routine least square analysis of the data would provide C, C^{-} and $C^{-}C$, the suggested use of the shorted matrix is more attractive compared to alternatives proposed earlier. It requires lesser number of additional matrix inversions, and as a by-product gives $H^{-}_{m(c)}$ which can be used to test the consistency of the equation (6.11).

Acknowledgement. The authors wish to thank Professor David H. Carlson for drawing their attention to the paper [4] by Ando.

REFERENCES

1. W.N. Anderson. Jr., Shorted operators. SIAM J. Appl. Math. 20:520-525 (1971).

2. W.N. Anderson, Jr. and R.J. Duffin, Series and parallel addition of matrices, J. Math. Anal. Appl. 11:576-594 (1969).

3. W.N. Anderson, Jr. and G. Trapp, Shorted operators II, SIAM J. Appl. Math. 28:60-71 (1975).

4. T. Ando, Generalized Schur Complements, Linear Algebra Appl. 27:173-186 (1979).

5. D. Carlson, Matrix decompositions involving the Schur complement, SIAM J. Appl. Math. 28:577-587 (1975).

6. W.G. Cochran and G.M. Cox, Experimental Designs (second edition) Wiley, New York, 1957.

7. R.J. Duffin and T.D. Morley, Almost definite operators and electromechanical systems, SIAM J. Appl. Math. 35: 21-30 (1978).

8. R.J. Duffin and G.E. Trapp, Hybrid addition of matrices - network theory concept, Applicable Analysis 2:241-254 (1972).

9. O. Kempthorne, The design and analysis of experiments, Wiley, New York, 1952.

10. M.G. Krein, The theory of self-adjoint extensions of semi-bounded Hermitian operators and its applications, Mat. Sbornik N.S. 20(62):431-495 and 21(63):365-404 (1947).

11. T.O. Lewis and T.G. Newman, Pseudoinverses of positive semi-definite matrices, SIAM J. Appl. Math. 10:701-703 (1968).

12. S.K. Mitra, Fixed rank solutions of linear matrix equations, Sankhyā, Ser A 35:387-392 (1972).

13. S.K. Mitra and M.L. Puri, On parallel sum and difference of matrices, J. Math. Anal. Appl. 44:92-97 (1973).

14. S.K. Mitra and M.L. Puri, Shorted operators and generalized inverses of matrices, Linear Algebra Appl. 25:45-56 (1979).

15. S.K. Mitra and C.R. Rao, Some results in estimation and tests of linear hypothesis under the Gauss-Markoff model, Sankhyā Ser. A, 30:281-290 (1968).

16. S.K. Mitra and C.R. Rao, Projections under seminorms and generalized Moore-Penrose inverses, Linear Algebra Appl. 4:155-167 (1974).

17. S.K. Mitra and G.E. Trapp, On hybrid addition of matrices, Linear Algebra Appl. 10:19-35 (1975).

18. C.R. Rao, Unified theory of linear estimation, Sankhyā, Ser. A. 33: 371-394 (1971).

19. C.R. Rao and S.K. Mitra, Generalized Inverse of Matrices and Its Applications, Wiley, New York, 1971.

END

DATE
FILMED

1 81

DTIC